Journal of Forensic and Legal Medicine 20 (2013) 1093–1097

Contents lists available at ScienceDirect

Journal of Forensic and Legal Medicine

journal homepage: www.elsevier.com/locate/jflm

Short report

Forensic and population genetic analyses of eighteen non-CODIS miniSTR loci in the Korean population

Han Jun Jin, Ph.D Professor [a], Ki Cheol Kim, MSc Ph.D Student [b], Cha Eun Yoon, MSc Researcher [c], Wook Kim, Ph.D Professor [b,*]

[a] Department of Nanobiomedical Science, Dankook University, Cheonan 330-714, Republic of Korea
[b] Departement of Biological Sciences, Dankook University, Cheonan 330-714, Republic of Korea
[c] Biowithus Life Science Institute, Seoul 133-120, Republic of Korea

ARTICLE INFO

Article history:
Received 16 April 2013
Received in revised form
23 September 2013
Accepted 28 September 2013
Available online 14 October 2013

Keywords:
miniSTR
Degraded DNA
Multiplex PCR
Forensic genetics
Koreans

ABSTRACT

We analyzed the variation of eighteen miniSTR loci in 411 randomly chosen individuals from Korea to increase the probability that a degraded sample can be typed, as well as to provide an expanded and reliable population database. Six multiplex PCR systems were developed (multiplex I: D1S1677, D2S441 and D4S2364; multiplex II: D10S1248, D14S1434 and D22S1045; multiplex III: D12S391, D16S3253 and D20S161; multiplex IV: D3S4529, D8S1115 and D18S853; multiplex V: D6S1017, D11S4463 and D17S1301; multiplex VI: D5S2500, D9S1122 and D21S1437). Allele frequencies and forensic parameters were calculated to evaluate the suitability and robustness of these non-CODIS miniSTR systems. No significant deviation from Hardy–Weinberg equilibrium expectations were observed, except for D4S2364, D5S2500 and D20S161 loci. A multidimensional scaling plot based on allele frequencies of the six miniSTR loci (D1S1677, D2S441, D4S2364, D10S1248, D14S1434 and D22S1045) showed that Koreans appeared to have most genetic affinity with Chinese and Japanese than to other Eurasian populations compared here. The combined probability of match calculated from the 18 miniSTR loci was 2.902×10^{-17}, indicating a high degree of polymorphism. Thus, the 18 miniSTR loci can be suitable for recovering useful information for analyzing degraded forensic casework samples and for adding supplementary genetic information for a variety of analyses involving closely related individuals where there is a need for additional genetic information.

1. Introduction

Autosomal short tandem repeat (STR) markers are among the most powerful tools for forensic genetics including individual identity and paternity testing because of their numerous alleles, high genetic diversity and stable heredity in the human genome.[1,2] However, many forensic laboratories often encounter difficulties in further analyzing a poor-quality sample. Exposure of DNA to elements or to fire for a certain length of time, degradation can start due to bacterial, biochemical, or oxidative processes.[3] Thus, forensic DNA samples that are often subjected to environmental factors promoting degradation can pose challenges for data interpretation. In such a circumstance, loss of a signal is typically observed with a larger-sized STR amplicon. Due to the extensive DNA fragmentation in many forensic cases, standard STR testing is often inadequate by using a commercially available multiplex STR genotyping kit. Another approach to recover information from degraded DNA samples is to use the reduced size of the PCR products. Reducing the size of the PCR product (miniSTR) by moving primers in as close as possible to the STR repeat region is an established and highly effective method for improving the forensic analysis of highly degraded DNA samples.

Applications of STR analysis in forensic casework benefit from large population databases for estimating the probability of identity by chance.[4,5] The use of additional STR markers would provide sufficient forensic parameters for more difficult cases in paternity or maternity analyses, such as deficient cases (i.e., only the alleged father and the child are included), missing persons, or in case of mutations. Thus, it is important that the STR databases continue to be expanded and made more reliable to provide a better tool for forensic analysis. National DNA databases initiated in Korea in 2010 comprise the 13 CODIS markers in the United States as well as the well-known D2S1338, D19S433, Penta E and Penta D loci. Many forensic communities have proposed the inclusion of additional

* Corresponding author. Tel.: +82 70 7154 3441; fax: +82 41 559 7861.
E-mail address: wookkim@dankook.ac.kr (W. Kim).

1752-928X/$ – see front matter
http://dx.doi.org/10.1016/j.jflm.2013.09.027

loci, since the potential false matches with a large number of comparisons being made within and between databases.[6–8] Thus, the analysis of extended non-CODIS STR markers may be a potentially powerful tool for forensic analyses in the Korean population.

In this study, we present the distribution of allelic frequencies and forensic efficiency parameters of 18 non-CODIS miniSTR loci (Supplementary Table S1; multiplex I: D1S1677, D2S441 and D4S2364; multiplex II: D10S1248, D14S1434 and D22S1045; multiplex III: D12S391, D16S3253 and D20S161; multiplex IV: D3S4529, D8S1115 and D18S853; multiplex V: D6S1017, D11S4463 and D17S1301; multiplex VI: D5S2500, D9S1122 and D21S1437) from 411 randomly selected individuals in Korea.

2. Materials and methods

2.1. Sample collection and DNA extraction

Blood or buccal swab samples were collected from 411 unrelated healthy individuals in Cheonan, Korea. The familial history of participants was obtained to exclude their relatives from the analysis. The surname was not considered, because, proportions of top five surnames were reached to 60% among more than 200 surnames in Korea.[9] A separate written informed consent was obtained from all the donors before collecting their blood or buccal swab. Genomic DNA was extracted from whole blood or buccal cells by the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). The DNA samples were quantified using the NanoDrop 2000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

2.2. PCR amplification

For genotyping of the 18 miniSTR loci (Supplementary Table S1), six triplex PCRs were performed using the primer sets mentioned in previous studies.[10,11] All these loci are located on different chromosomes and D2S441 and D22S1045 are included in the Extended European Standard Set (ESS). Several fluorescent dyes for primer labeling were modified to accommodate the subsequent PCR fragment detection in capillary electrophoresis. The forward primers for D4S2364, D10S1248, D16S3253, D8S1115, D6S1017 and D5S2500 were labeled with FAM. The primers for D2S441, D14S1434, D20S161, D18S853, D11S4463 and D9S1122 were labeled with HEX. The primers for D1S1677, D22S1045, D12S391, D3S4529, D17S1301 and D21S1437 were labeled with NED (Supplementary Table S2). All primer sets were designed to amplify the PCR products with a maximum fragment size of <145 bp. Each multiplex PCR reaction was performed using a total volume of 10 μl containing 1 ng genomic DNA, 1 × PCR Gold buffer (Applied Biosystems, Foster City, CA, USA), 1.0 U AmpliTaq Gold® DNA polymerase (Applied Biosystems), 1.5 mM $MgCl_2$, 200 μM of each dNTP, and a primer concentration adjusted to balance the peak height. Thermal cycling was conducted on the GeneAmp® PCR System 9700 (Applied Biosystems) with slight modification in the PCR condition.[12] Initial denaturation was performed at 95 °C for 10 min, followed by 30 cycles of 94 °C for 30 s, 55 °C (except for multiplex III: 58 °C; multiplex IV: 64 °C; and multiplex V, and VI: 61 °C) for 30 s, 72 °C for 45 s, and a final extension of 60 °C for 45 min.

2.3. Genotyping

For the genotyping, the PCR products were mixed with the GeneScan™ 500 ROX™ Size Standard (Applied Biosystems). Capillary electrophoresis was performed on the ABI PRISM™ 310 Genetic Analyzer (Applied Biosystems). The results were analyzed, and allele designations were determined by the GeneMapper 4.1 software (Applied Biosystems) with sequenced allelic ladders. The allelic ladders for multiplexes I–VI were created using a combination of individual templates, which represent the range for alleles observed by Han et al.[12] and in this study. At least two different homozygous samples for each miniSTR marker were sequenced to confirm the allele nomenclature by a procedure described previously.[13] Commercial DNA standard 9947A and 9948 (Promega, Madison, WI, USA) were genotyped as the positive standard reference for quality control. In addition, ddH_2O was used as the negative control. A concordant study was performed to ensure reproducible and accurate genotyping by a re-genotyping procedure by using more than 20 samples for each multiplex.

2.4. Statistical analyses

General goodness-of-fit test of x^2-test were used to assess Hardy–Weinberg equilibrium (HWE). Possible divergence from HWE was also examined by the Exact Test.[14] These statistical analyses were performed with the PowerMarker v3.25[15] and Arlequin v3.5.[16] The genetic distance was estimated by shared allele genetic distance (D_{SA}).[17] The genetic distance values were also displayed as a multidimensional scaling (MDS) plot by using SPSS statistics 19 (SPSS Korea, Seoul, Korea). Important parameters of forensic paternity testing were calculated using PowerStats v1.2.[18]

3. Results and discussion

We analyzed six miniSTR multiplex systems for the typing of 18 non-CODIS loci. None were located on the same chromosome. We assessed the statistical parameters for applications to forensic and population genetic studies by using these 18 miniSTR loci from the Korean samples surveyed in this study. The distributions of allelic frequencies and forensic parameters of the 18 minSTR loci from the 411 Koreans are summarized in Table 1. The exact tests were performed on all 18 miniSTR loci to determine HWE and the p-values are listed in Table 1. No significant deviation from the HWE expectations was observed, except for D4S2364, D5S2500 and D20S161 loci. After Bonferroni correction, these three loci were still significant in the HWE test ($p < 0.00278$). Deviation from HWE of D4S2364 in Korean subjects and D5S2500 in Chinese subjects has been previously reported,[12,13,19] but deviation of the D20S161 locus has not been reported. Additional studies with larger sample sizes are necessary to understand the cause of deviation from the equilibrium model at the three loci surveyed in this study.

Based on our result of allele distributions of 18 STR loci, the number of alleles varied from four at D4S2364 to 12 at D8S1115 (Supplementary Table S1). The observed heterozygosity (H_{obs}) was greater than 0.7 for 12 out of 18 loci. Of the 18 loci, D11S4463 was the most informative miniSTR locus (H_{obs} value at 0.7835), whereas D5S2500 was the least informative miniSTR locus (H_{obs} value at 0.6010). Individual probability of match (PM) values ranged from 0.0597 at D12S391 to 0.2318 at D4S2364. The power of discrimination (PD) varied from 0.7682 at D4S2364 to 0.9403 at D12S391. The paternity power of exclusion (PPE) ranged from 0.3682 at D2S441 to 0.5677 at D11S4463. The combined PM calculated from the 18 minSTR loci studied was 2.902×10^{-17}, and the PPE was estimated to be 0.9999871. In contrast, Coble et al. reported two miniplex systems called NC01 (D10S1248, D14S1434 and D22S1045) and NC02 (D1S1677, D2S441 and D4S2364), showing that combined PM and PPE from the six miniSTR loci were 7.409×10^{-6} and 0.974, respectively.[20] The authors proposed the inclusion of further or different STR loci to satisfy the increasing requirements of forensic parameters for obtaining additional information. Thus, our results indicate that these 18 miniSTR loci have relatively higher probability of exclusion as well as PD and therefore can be useful in forensic analysis.

Table 1
Allele frequencies and statistical parameters for eighteen miniSTR loci in the Korean population ($n = 411$).

Allele	D1S1677	D2S441	D3S4529	D4S2364	D5S2500	D6S1017	D8S1115	D9S1122	D10S1248
6									
7				0.034					
8				0.174		0.268	0.001	0.007	
9		0.001		0.470			0.173	0.004	
10	0.007	0.181		0.322		0.328	0.001	0.091	
11	0.002	0.466				0.026	0.006	0.147	0.004
11.3		0.017							
12	0.007	0.170	0.004		0.002	0.296		0.303	0.072
13	0.120	0.056	0.165			0.073	0.001	0.369	0.387
14	0.467	0.100	0.302		0.370	0.007	0.023	0.067	0.238
15	0.331	0.009	0.322		0.002	0.002	0.038	0.012	0.206
16	0.058		0.165		0.001		0.477		0.079
17	0.006		0.041		0.345		0.217		0.015
18					0.251		0.050		
19					0.001		0.012		
20					0.024		0.001		
21									
22									
23					0.002				
24									
25									
H_{obs}[a]	0.6618	0.6594	0.7348	0.7202	0.6010	0.7518	0.6618	0.7129	0.7664
H_{exp}[b]	0.6542	0.7076	0.7486	0.6441	0.6804	0.7270	0.6912	0.7377	0.7396
X^2-test (P)	1.0000	0.0006	0.5333	0.0001	0.0000	0.8019	0.9991	0.0001	0.6280
Exact test (P)[c]	0.9267	0.0682	0.5673	0.0000	0.0000	0.5470	0.4185	0.2282	0.5482
PM[d]	0.1845	0.1229	0.1063	0.2318	0.1616	0.1324	0.1376	0.1125	0.1104
PD[e]	0.8155	0.8771	0.8937	0.7682	0.8384	0.8676	0.8624	0.8875	0.8896
PIC[f]	0.5945	0.6712	0.7059	0.5778	0.6112	0.6771	0.6501	0.6963	0.6992
PPE[g]	0.3717	0.3682	0.4841	0.4602	0.2885	0.5119	0.3705	0.4485	0.5383

Allele	D11S4463	D12S391	D14S1434	D16S3253	D17S1301	D18S853	D20S161	D21S1437	D22S1045
6				0.068					
7				0.001	0.007				
8				0.146	0.007			0.001	
9	0.007			0.060	0.040				
10	0.001		0.120	0.336	0.079	0.002		0.089	
11	0.004		0.153	0.353	0.189	0.376		0.125	0.190
11.3									
12	0.049		0.030	0.012	0.405	0.062		0.019	0.002
13	0.195		0.243	0.012	0.213	0.264		0.039	0.001
14	0.314		0.436	0.011	0.049	0.218		0.524	0.005
15	0.279	0.019	0.013		0.011	0.075	0.045	0.161	0.299
16	0.124	0.007	0.004	0.001			0.162	0.035	0.251
17	0.027	0.102					0.412	0.006	0.227
18	0.001	0.297				0.002	0.155		0.021
19		0.240					0.135		0.004
20		0.153					0.082		
21		0.090					0.009		
22		0.051					0.001		
23		0.026							
24		0.013							
25		0.001							
H_{obs}[a]	0.7835	0.7421	0.6837	0.7202	0.7762	0.7324	0.7518	0.6740	0.7810
H_{exp}[b]	0.7674	0.8085	0.7120	0.7329	0.7445	0.7320	0.7529	0.6725	0.7594
x^2-test (P)	0.1527	0.5570	0.9168	0.8198	0.9032	0.9431	0.0036	0.3587	0.5329
Exact test (P)[c]	0.1113	0.3350	0.7806	0.1822	0.7000	0.7811	0.0006	0.3368	0.0653
PM[d]	0.0951	0.0597	0.1223	0.1189	0.1063	0.1143	0.0971	0.1397	0.1044
PD[e]	0.9049	0.9403	0.8777	0.8811	0.8937	0.8857	0.9029	0.8603	0.8956
PIC[f]	0.7294	0.7837	0.6693	0.6902	0.7096	0.6873	0.7224	0.6403	0.7174
PPE[g]	0.5677	0.4963	0.4035	0.4602	0.5556	0.4801	0.5129	0.3880	0.5643

[a] H_{obs}: observed heterozygosity.
[b] H_{exp}: expected heterozygosity.
[c] Exact test (Monte Carlo method).
[d] PM, probability of match.
[e] PD, power of discrimination.
[f] PIC, polymorphism information content.
[g] PPE, paternity power of exclusion.

A population comparison based on the D_{SA} values calculated from the allele frequencies of the six miniSTR loci in NC01 and NC02 systems between the Korean population and other published populations is shown in Supplementary Table S3.[21–30] An MDS plot of D_{SA} based on the results presented in Supplementary Table S3 is depicted in Fig. 1. The plot shows two distinct clusters (East Asians and African/Europeans/Americans), one intermediate population (Indian–Singaporean) and one out-population (Brazilian Mulatto).

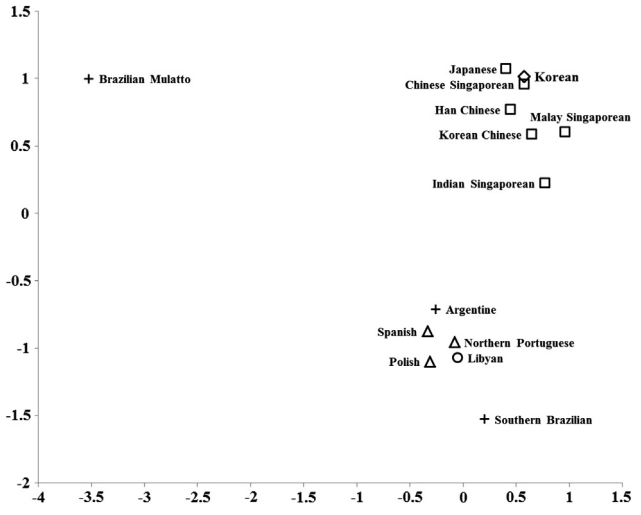

Fig. 1. Multidimensional scaling plot of shared allele genetic distance (D_{SA}) based on the results presented in Supplementary Table S3 (stress = 0.17). Korean population is represented by diamond, Asians by squares, Europeans by triangles, Americans by crosses, and Africans by circles.

As expected, Koreans appeared to have the closest genetic affinity with Chinese and Japanese (Fig. 1). This result was consistent with a previous report derived from the datasets of mitochondrial DNA and Y chromosome markers.[31] Indian Singaporeans were located as an intermediate between the East Asian cluster and others. It has been reported that peopling of the Indian subcontinent is characterized by a complex history, with contributions from different ancestral populations.[32] The Brazilian Mulatto population was isolated from the main cluster of the MDS plot. This may reflect their mixed ethnicity, because Mulattos are descendants of European and African ancestry/lineage.[29]

In summary, our study demonstrates that the 18 non-CODIS minSTR loci are highly polymorphic. Additional these miniSTR markers, together with the conventional STRs of CODIS, multi-locus commercial kits (AmpFlSTR® Identifiler® and Powerplex® 16) or miniSTR commercial kit (AmpFlSTR® Minifiler®) may help resolve more complex forensic cases. Thus, the 18 miniSTR markers can serve as useful complements to other commercial STR systems for adding supplementary genetic information not only for the typing of degraded samples but also for a variety of analyses involving closely related individuals when additional genetic information is required. Further surveys of the miniSTR systems to obtain additional evidence-based data would provide further insights.

Ethical approval

None.

Funding

National Research Institute of Cultural Heritage of Cultural Heritage Administration, Republic of Korea.

Conflict of interest

None declared.

Acknowledgments

We would like to thank all volunteers for providing DNA samples. Special thanks go to S.S. Hong for technical assistance of this survey. This study, which forms a part of the project (NRICH-1017-B07F), has been achieved with the support of Conservation Technology Research and Development project, which has been hosted by National Research Institute of Cultural Heritage of Cultural Heritage Administration, Republic of Korea.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jflm.2013.09.027.

References

1. Butler JM. Short tandem repeat typing technologies used in human identity testing. *Biotechniques* 2007;**43**:2–5.
2. Gill P, Werrett DJ, Budowle B, Guerrieri R. An assessment of whether SNPs will replace STRs in national DNA databases–joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDAM). *Sci Justice* 2004;**44**:51–3.
3. Bär W, Kratzer A, Mächler M, Schmid W. Postmortem stability of DNA. *Forensic Sci Int* 1988;**39**:59–70.
4. Lessig R, Willuweit S, Krawczak M, Wu FC, Pu CE, Kim W, et al. Asian online Y-STR haplotype reference database. *Leg Med (Tokyo)* 2003;**5**(Suppl. 1):S160–3.
5. Willuweit S, Roewer L. Y chromosome haplotype reference database (YHRD): update. *Forensic Sci Int Genet* 2007;**1**:83–7.
6. Weir BS. The rarity of DNA profiles. *Ann Appl Stat* 2007;**1**:358–70.
7. Schneider PM. Expansion of the European standard set of DNA database loci—the current situation. *Profiles DNA* 2009;**12**:6–7.
8. Hill CR, Duewer DL, Kline MC, Sprecher CJ, McLaren RS, Rabbach DR, et al. Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex® ESX 17 and ESI 17 Systems. *Forensic Sci Int Genet* 2011;**5**:269–75.
9. Park SW, Hwang CH, Cho EM, Park JH, Choi BO, Chung KW. Development of a Y-STR 12-plex PCR system and haplotype analysis in a Korean population. *J Genet* 2009;**88**:353–8.
10. Hill CR, Kline MC, Coble MD, Butler JM. Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J Forensic Sci* 2008;**53**:73–80.
11. Asamura H, Fujimori S, Ota M, Fukushima H. MiniSTR multiplex systems based on non-CODIS loci for analysis of degraded DNA samples. *Forensic Sci Int* 2007;**173**:7–15.
12. Han MS, Kim YS, Jin HJ, Kim JJ, Kwak KD, Lee JE, et al. Forensic genetic analysis of nine miniSTR loci in the Korean population. *Leg Med (Tokyo)* 2009;**11**:209–12.
13. Chung U, Shin KJ, Park MJ, Kim NY, Yang WI, Cho SH, et al. Population data of nine miniSTR loci in Koreans. *Forensic Sci Int* 2007;**168**:e51–3.
14. Guo SW, Thompson EA. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* 1992;**48**:361–72.
15. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 2005;**21**:2128–9.
16. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* 2005;**1**:47–50.
17. Chakraborty R, Jin L. A unified approach to study hypervariable polymorphisms: statistical considerations of determining relatedness and population distances. *EXS* 1993;**67**:153.
18. Tereba A. Tools for analysis of population statistics. *Profiles DNA* 1999;**2**:14–6.
19. Zhu B, Shen C, Wang H, Yang G, Yan J, Qin H, et al. Genetic diversities of 21 non-CODIS autosomal STRs of a Chinese Tibetan ethnic minority group in Lhasa. *Int J Legal Med* 2011;**125**:581–5.
20. Coble MD, Butler JM. Characterization of new miniSTR loci to aid analysis of degraded DNA. *J Forensic Sci* 2005;**50**:43–53.
21. Asamura H, Uchida R, Takayanagi K, Ota M, Fukushima H. Allele frequencies of the six miniSTR loci in a population from Japan. *Int J Legal Med* 2006;**120**:182–4.
22. Bai R, Shi M, Yu X, Lv J, Tu Y. Allele frequencies for six miniSTR loci of two ethnic populations in China. *Forensic Sci Int* 2007;**168**:e25–8.
23. Yong R, Gan L, Coble M, Yap E. Allele frequencies of six miniSTR loci of three ethnic populations in Singapore. *Forensic Sci Int* 2007;**166**:240–3.
24. Martín P, García O, Albarrán C, García P, Yurrebaso I, Alonso A. Allele frequencies of six miniSTR loci (D10S1248, D14S1434, D22S1045, D4S2364, D2S441 and D1S1677) in a Spanish population. *Forensic Sci Int* 2007;**169**:252–4.
25. Reichert M, Pawłowski R. Population genetics of six miniSTR loci (D1S1677, D2S441, D4S2364, D10S1248, D14S1434, D22S1045) in a Polish population. *Leg Med (Tokyo)* 2009;**11**:147–8.
26. Lagoa AM, Martins TV, Cainé LM, Pinheiro MF. Allele frequencies of six miniSTR loci in the population of Northern Portugal. *Forensic Sci Int Genet* 2008;**2**:379–81.
27. Vullo C, Borosky A, Romanini C, Catelli L, Yamamoto T. Frequency data for 12 mini STR loci in Argentina. *Forensic Sci Int Genet* 2010;**4**:e79–81.
28. Raimann PE, de Oliveira A, Rodenbusch R, Picanço JB, Albuquerque T, Alho CS. Genetic data for D1S1677, D2S441, D4S2364, D10S1248, D14S1434 and D22S1045 miniSTR loci from the state of Rio Grande do Sul, Southern Brazil. *Forensic Sci Int Genet* 2012;**6**:e42.

29. Aranda XG, Lage CAC, Planz JV, Eisenberg AJ, Moura-Neto RS, Silva R. Genetic composition of six miniSTR in a Brazilian Mulatto sample population. *J Forensic Leg Med* 2011;**18**:184–6.
30. Aranda XG, Moura-Neto RS, Al-Deib AWA, Aboud AI, Planz JV, Eisenberg AJ, et al. Genetic data for D1S1677, D2S441, D4S2364, D10S1248, D14S1434 and D22S1045 miniSTR loci from Libya. *Forensic Sci Int Genet* 2010;**4**:267–8.
31. Jin HJ, Tyler-Smith C, Kim W. The peopling of Korea revealed by analyses of mitochondrial DNA and Y-chromosomal markers. *PloS ONE* 2009;**4**:e4210.
32. Xing J, Watkins WS, Hu Y, Huff CD, Sabo A, Muzny DM, et al. Genetic diversity in India and the inference of Eurasian population expansion. *Genome Biol* 2010;**11**:R113.